ORIGINAL REPORTS

Early Performance on an Eye Surgery Simulator Predicts Subsequent Resident Surgical Performance

Ramak Roohipoor, MD,*,† Mehdi Yaseri, PhD, MPH,†,‡ Amir Teymourpour, MSc,‡ Carolyn Kloek, MD,* John B. Miller, MD,* and John I. Loewenstein, MD*

*Department of Ophthalmology, Massachusetts Eye and Ear, Harvard Medical School, Boston, Massachusetts; †Eye Research Center, Farabi Eye Hospital, Tehran University of Medical Sciences, Tehran, Iran; and ‡Biostatistic Department, Tehran University of Medical Sciences, Tehran, Iran

OBJECTIVE: To examine early performance on an eye surgery simulator and its relationship to subsequent live surgical performance in a single large residency program.

DESIGN: Retrospective study.

SETTING: Massachusetts Eye and Ear, Harvard Medical School, Department of Ophthalmology.

METHODS: In a retrospective study, we compared performance of 30 first-year ophthalmology residents on an eye surgery simulator to their surgical skills as third-year residents. Variables collected from the eye surgery simulator included scores on the following modules of the simulator (Eyesi, VRmagic, Mannheim, Germany): antitremor training level 1, bimanual training level 1, capsulorhexis level 1 (configured), forceps training level 1, and navigation training level 1. Subsequent surgical performance was assessed using the total number of phacoemulsification cataract surgery cases for each resident, as well as the number performed as surgeon during residency and scores on global rating assessment of skills in intraocular surgery (GRASIS) scales during the third year of residency. Spearman correlation coefficients were calculated between each of the simulator performance and subsequent surgical performance variables. We also compared variables in a small group of residents who needed extra help in learning cataract surgery to the other residents in the study.

MAIN OUTCOME MEASURES: Relationships between Eyesi scores early in residency and surgical performance measures in the final year of residency.

RESULTS: A total of 30 residents had Eyesi data from their first year of residency and had already graduated so that all subsequent surgical performance data were available. There was a significant correlation between capsulorhexis task score on the simulator and total surgeries (r = 0.745, p = 0.008). There was a significant correlation between antitremor training level 1 (r = 0.554, p = 0.040), and forceps training level 1 (r = 0.622, p = 0.023) with primary surgery numbers.
There was a significant correlation between forceps training level 1 (r = 0.811, p = 0.002), and navigation training level 1 (r = 0.692, p =0.013) with total GRASIS score. There was a significant inverse correlation between total GRASIS score and residents in need of extra help (r = −0.358, p =0.003).

CONCLUSION: Module scores on an eye surgery simulator early in residency may predict a resident's future performance in the operating room. These scores may allow early identification of residents in need of supplemental training in cataract surgery. (J Surg Ed ■:■■■-■■■. © 2017 Association of Program Directors in Surgery. Published by Elsevier Inc. All rights reserved.)

KEY WORDS: ophthalmology education, applicant evaluation, surgical performance, Eyesi, simulator

COMPETENCIES: Patient Care

INTRODUCTION

Surgical simulators have been widely used for training in neurosurgery, gastroentrology, laparoscopic surgery, orthopedics, and ophthalmology.[1-6] There is increasing evidence

☆This article has been presented in part at American Academy of Ophthalmology annual meeting in 2015.

☆☆Financial support: None.

★Financial disclosure: No author has a financial or proprietary interest in any material or method mentioned.

Correspondence: Inquiries to John I. Loewenstein, MD, Retina Service, Department of Ophthalmology, Massachusetts Eye and Ear, Harvard Medical School, 243 Charles Street, Boston, MA 02114; e-mail: John_Loewenstein@meei.harvard.edu

1931-7204/$30.00
http://dx.doi.org/10.1016/j.jsurg.2017.04.002

that surgical simulator training improves residents' surgical performance in the operating room (OR) and improves surgical outcomes.[5] It is difficult to predict the potential for the acquisition of surgical skills in a novice surgeon. Although most residents can be adequately trained surgically, the occasional resident may have difficulty achieving surgical competencies within the timeframe of a residency program. In these unfortunate situations, it is not unusual to discover these surgical deficiencies late in residency. Many ophthalmology surgical programs have residents perform most of their surgery in their final year.

The Eyesi simulator has been used to train residents in cataract surgery at many ophthalmology programs including the Harvard Medical School ophthalmology residency. It can generate scores for performance in multiple modules based upon errors made, path length, time taken for different simulator tasks, and other variables. Construct validity for various modules of the Eyesi eye surgery simulator (VRmagic, Mannheim, Germany) has been demonstrated by several authors.[7-10]

To date, there are few formal studies that explore the use of early assessment tools to predict future surgical performance.[11-13] Some general surgeons have proposed that the incorporation of technical proficiency skills, such as visual spatial perception, in residency selection may help identify those candidates with an aptitude for a surgical specialty.[14,15] Although simulators for teaching and assessing surgical skills are become more widely accepted in ophthalmology, there is insufficient evidence as to whether early simulator performance can predict subsequent intraoperative surgical performance.

In this study, we hope to address this gap in the literature, hypothesizing that residents' early simulator performance may predict future surgical skills in the OR. A better prediction method would allow training programs to provide early intervention for residents who will need more teaching and practice than the standard curriculum provides.

METHODS

A retrospective review of Harvard Medical School ophthalmology residents' performance from July 2011 to July 2014 on the Eyesi simulator was performed. The Massachusetts Eye and Ear Institutional Review Board determined that this study was exempt as an educational study. There were 32 eligible residents. One was excluded because of prior experience in ophthalmology residency outside the United States, and the other was because of insufficient information on the Eyesi. We considered early attempts by each resident during the first 3 months of residency in each of the following tasks from the simulator: antitremor training level 1, bimanual training level 1, capsulorhexis level 1 (configured), forceps training level 1, and navigation training level 1. We excluded the first 3 attempts to minimize familiarity with the Eyesi as a variable, reviewing the next 8 attempts for each resident. Eyesi scores were calculated by the trainer software, based on variables such as time, tissue injury, microscope focus, path length, and others. We excluded scores of 0 with instruments in the eye for less than 10 seconds.

Intraoperative performance of residents' cataract surgery was assessed in the following 2 ways:

(1) Global rating assessment of skills in intraocular surgery (GRASIS) forms. In our residency program, faculty complete a validated performance scale called GRASIS[15] during the third-year resident's cataract surgery rotations. We considered the following items from the GRASIS: instrument handling, flow of operation, time and motion, treatment of ocular structures and other tissues, use of nondominant hand, and overall performance. These measures were graded on a modified Likert scale from 1 to 5.

(2) Number of cataract surgeries performed by a resident during all of residency training. We assessed both the total number of surgeries in which a resident participated, and the number of surgeries as primary surgeon. Although there are many factors that influence the number of cases a resident performs as primary surgeon, attending confidence in the resident's surgical skill is a factor in the number of cases in which the resident acts as primary surgeon.

In addition, there was a group of residents whom the residency program director identified as needing extra help in learning cataract surgery based on feedback from cataract surgery faculty. We compared Eyesi and subsequent surgical performance in this group to these parameters in the remaining group of residents.

Statistical Analysis

We eliminated the first 3 attempts in each task on the Eyesi for each resident during the first year of training and we considered the next 8 attempts during the first 3 months of the start of residency. To obtain the mean improvement of scores after each attempt, we calculated the regression coefficient of score versus the attempt number for each task of each resident in early performance evaluation in a linear mixed model. The mean total task score of each resident was calculated for each task during the resident's early performance. Based on these findings, we obtained the Spearman correlation of early simulator performance, total surgeries, total primary surgeries, and mean GRASIS score for each resident. We also determined the Spearman correlation of these variables with the need for extra help in learning cataract surgery. Subsequently, we calculated the

TABLE 1. Early Scores of Residents on the Eyesi Simulator

	Antitremor Training Level 1	Bimanual Training Level 1	Capsulorhexis Level 1 (Configured)	Forceps Training Level 1	Navigation Training Level 1	Total
Mean early task score	59.3 ± 21.9	83.4 ± 11.3	59.7 ± 12.3	69.7 ± 10.9	64.5 ± 13.4	66.9 ± 16.3
Improvement	9.2 ± 16	14.1 ± 8.2	0 ± 1.4	25 ± 25.8	3.4 ± 5.1	9.4 ± 15.6
Last early task score	71.4 ± 29.1	90.2 ± 6	64.9 ± 33.8	87.8 ± 11.7	54.3 ± 34.3	73.7 ± 27.1

mean change in primary surgeries and mean GRASIS score for each 10-step increase of mean improvement of scores after each attempt on Eyesi and mean total task score within each task using linear regression. Based on a linear mixed model, the intracluster correlation (ICC) of the raters was analyzed. ICC was interpreted as follows: 0 to 0.2 = poor agreement, 0.3 to 0.4 = fair agreement, 0.5 to 0.6 = moderate agreement, 0.7 to 0.8 = strong agreement, and >0.8 = almost perfect agreement.[16]

We performed all statistics using R software (R Core Team (2014). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL: http://www.R-project.org/).

RESULTS

The mean (± standard deviation) score for antitremor level 1, bimanual training level 1, capsulorhexis level 1, forceps training level 1, and navigation training level 1 was 59 ± 21.9, 83.4 ± 11.3, 59.7 ± 12.3, 69.7 ± 10.9, and 64.5 ± 13.4, respectively (Table 1). The mean (±standard deviation) GRASIS score and number of primary surgeries was 4.13 (±0.74) with median 4 (range: 3-5) and 143 (±25.5) with median 144 (range: 101-215) (Table 2).

There was a significant correlation between the capsulorhexis task score on the simulator and the total number of cataract surgeries (r = 0.74, p = 0.008). There was a significant correlation between antitremor training level 1 and forceps training level 1 with primary surgery numbers (Spearman correlation coefficient = 0.55; p = 0.04 and 0.62, p = 0.02, respectively). There was a significant correlation between forceps training level 1 and navigation training level 1 with total GRASIS score (r = 0.81, p = 0.002 and r = 0.69, p = 0.013, respectively). There was a significant correlation among total sum score for all above mentioned tasks (5 tasks) on Eyesi with total mean of GRASIS score (r = 0.368, p = 0.006) (Table 3 and Fig. 1) and also with total (r = 0.28, p = 0.021) and primary number (r = 0.44, p = 0.001) of cataract surgeries done by residents (Table 3; Figs. 2 and 3). Correlation between final task score and GRASIS score is shown in Table 3.

Six residents needing extra help in cataract surgery were identified. The mean difference in each task score among residents in need of extra help and the residents who did not need help is shown in Figure 4. There was a significant inverse correlation between total GRASIS score and residents in need of extra help (r = −0.358, p = 0.003; Table 3).

The mean improvement in each task on Eyesi is shown in Table 1. Improvement in antitremor training level 1 (r = 0.662, p = 0.019) and navigation training level 1 (r = 0.665, p = 0.026) scores on Eyesi were significantly correlated with primary surgeries (Table 3). Improvement in Eyesi capsulorhexis and navigation training level 1 score was very minimal in all residents (Fig. 5). Mean change in

TABLE 2. Surgical Performance Characteristics of Residents

	G* Score						
	Instrument Handling	Treatment of Ocular Structure	Time and Motion	Use of Nondominant Hand	Overall Performance	Independent Sx† Number	Total Sx† Number
Mean	4.1	4.1	3.9	4.2	4.1	143.0	306.8
SD	0.8	0.8	0.8	0.7	0.9	25.5	57.6
Min	2	2	2	3	2	101	174
Q1	3.125	3.5	3	4	3	125.25	285
Median	4	4	4	4	4	144	296
Q3	5	5	5	5	5	158.5	328.75
Max	5	5	5	5	5	215	417

*GRASIS.
†Surgery.

TABLE 3. Spearman Correlation Coefficient Between Early Performance on Eyesi and Operating Room Performance

		Total Surgery	Independent Surgery	Help	G Score Instrument Handling	G Score Treatment Ocular Structure	G Score Time Motion	G Score Use Nondominant Hand	G Score Overall Performance	Sum G Score
Score										
Antitremor training level 1	*r*	0.349	**0.554***	−0.091	0.339	0.339	0.288	0.049	0.121	0.280
	p	0.221	0.040	0.748	0.282	0.282	0.390	0.879	0.723	0.379
Bimanual training level 1	*r*	−0.046	0.107	−0.464	0.474	0.474	0.093	0.214	0.087	0.407
	p	0.888	0.740	0.110	0.166	0.166	0.812	0.553	0.823	0.243
Capsulorhexis level 1 (configured)	*r*	**0.745****	0.588	−0.518	0.061	−0.128	0.167	−0.264	−0.068	0.006
	p	0.008	0.057	0.084	0.868	0.724	0.667	0.461	0.862	0.987
Forceps training level 1	*r*	0.187	**0.622***	-0.448	**0.795****	**0.795****	0.607	**0.753****	**0.806****	**0.811****
	p	0.541	0.023	0.108	0.003	0.003	0.063	0.007	0.005	0.002
Navigation training level 1	*r*	0.204	0.499	−0.454	**0.646***	**0.646***	0.590	**0.723****	0.321	**0.692***
	p	0.483	0.069	0.089	0.023	0.023	0.056	0.008	0.336	0.013
All scores	*r*	**0.288***	**0.442****	**−0.358****	**0.377****	**0.353****	**0.326***	0.260	0.188	**0.368****
	p	0.021	0.000	0.003	0.005	0.008	0.021	0.055	0.191	0.006
Improvement										
Antitremor training level 1	*r*	0.531	**0.662***	0.000	0.126	0.126	−0.168	0.142	−0.049	0.133
	p	0.075	0.019	1.000	0.748	0.748	0.691	0.715	0.909	0.732
Bimanual training level 1	*r*	−0.108	−0.228	0.275	−0.162	−0.162	−0.132	0.464	0.088	−0.162
	p	0.799	0.588	0.474	0.728	0.728	0.803	0.295	0.868	0.728
Capsulorhexis level 1 (configured)	*r*	−0.164	0.269	0.194	−0.201	0.000	−0.259	−0.252	−0.203	−0.212
	p	0.631	0.424	0.545	0.578	1.000	0.500	0.483	0.600	0.556
Forceps training level 1	*r*	−0.109	−0.169	−0.219	0.548	0.548	0.460	0.225	0.426	0.511
	p	0.750	0.620	0.495	0.101	0.101	0.213	0.531	0.253	0.132
Navigation training level 1	*r*	0.518	**0.665***	−0.065	−0.203	−0.203	−0.072	−0.017	−0.268	−0.083
	p	0.102	0.026	0.841	0.600	0.600	0.866	0.966	0.520	0.831

All scores	*r*	0.125	0.140	-0.031	0.041	0.020	0.098	0.060	0.071	0.056
	p	0.374	0.319	0.821	0.791	0.898	0.547	0.698	0.664	0.715
Final score in early Eyesi performance										
Antitremor training level 1	*r*	0.226	**0.355**	-0.054	0.306	0.351	0.284	0.096	0.224	0.269
	p	0.438	0.212	0.847	0.333	0.263	0.397	0.766	0.508	0.398
Bimanual training level 1	*r*	−0.391	−**0.596***	−0.112	−0.089	−0.140	−0.346	−0.167	−0.123	−0.063
	p	0.209	0.041	0.715	0.808	0.699	0.362	0.644	0.752	0.862
Capsulorhexis level 1 (configured)	*r*	−0.593	−0.492	0.174	−0.015	0.084	−0.277	0.065	0.040	−0.016
	p	0.055	0.124	0.589	0.968	0.818	0.470	0.858	0.918	0.964
Forceps training level 1	*r*	0.423	0.319	−0.266	0.437	0.487	0.320	0.232	0.314	0.384
	p	0.150	0.288	0.358	0.179	0.129	0.367	0.492	0.376	0.244
Navigation training level 1	*r*	−0.417	−0.275	−0.181	0.151	0.107	0.346	−0.044	0.429	−0.046
	p	0.138	0.341	0.519	0.639	0.740	0.297	0.892	0.187	0.888
All scores	*r*	−0.200	−0.138	−0.068	0.123	0.138	0.118	0.013	0.204	0.054
	p	0.083	0.233	0.544	0.328	0.273	0.374	0.916	0.120	0.668

Bold entries are statistically significant at $p < 0.05$ or $p < 0.01$.
*indicates $p < 0.05$.
**indicates $p < 0.01$.

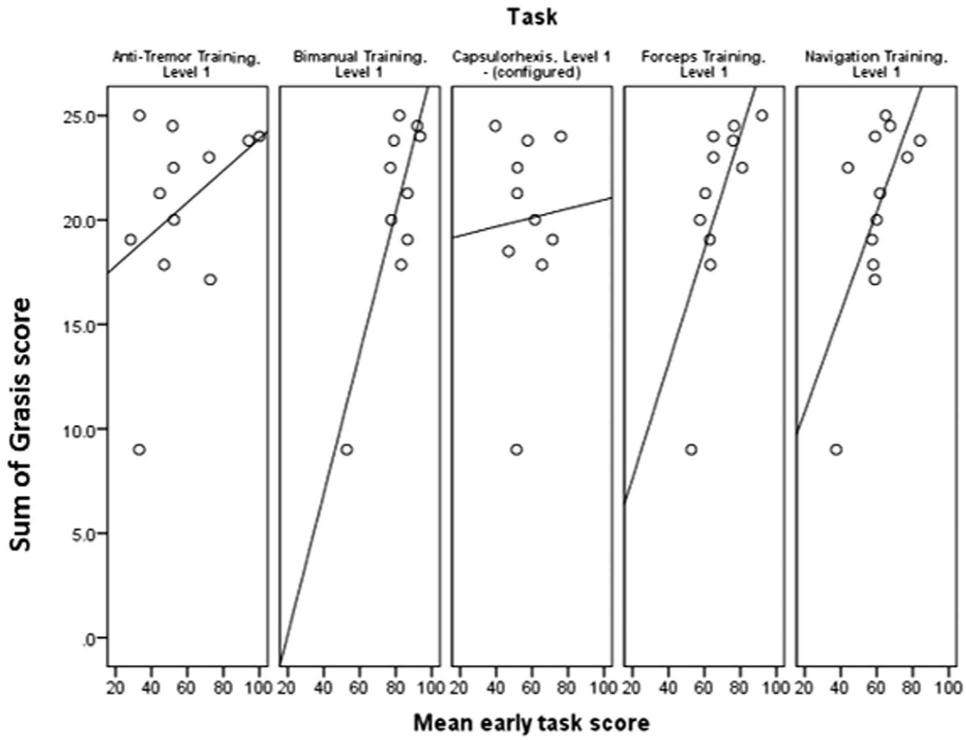

FIGURE 1. Scatter plot showing relationship between sum of GRASIS score and each early mean task score on Eyesi.

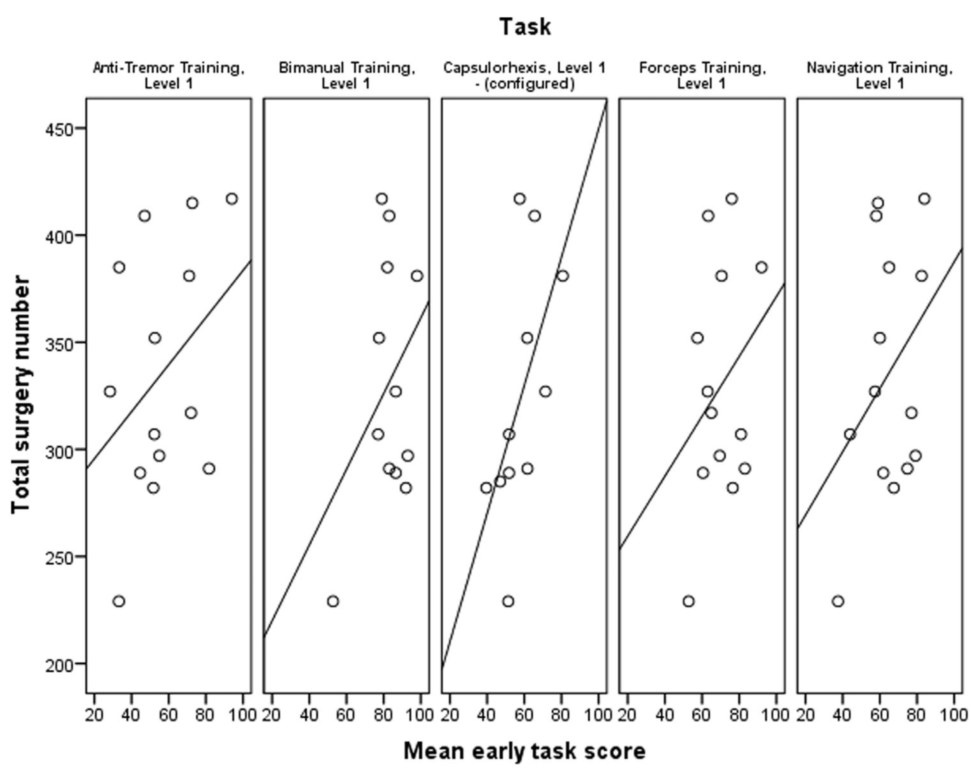

FIGURE 2. Scatter plot showing relationship between total number of cataract surgery and each early mean task score.

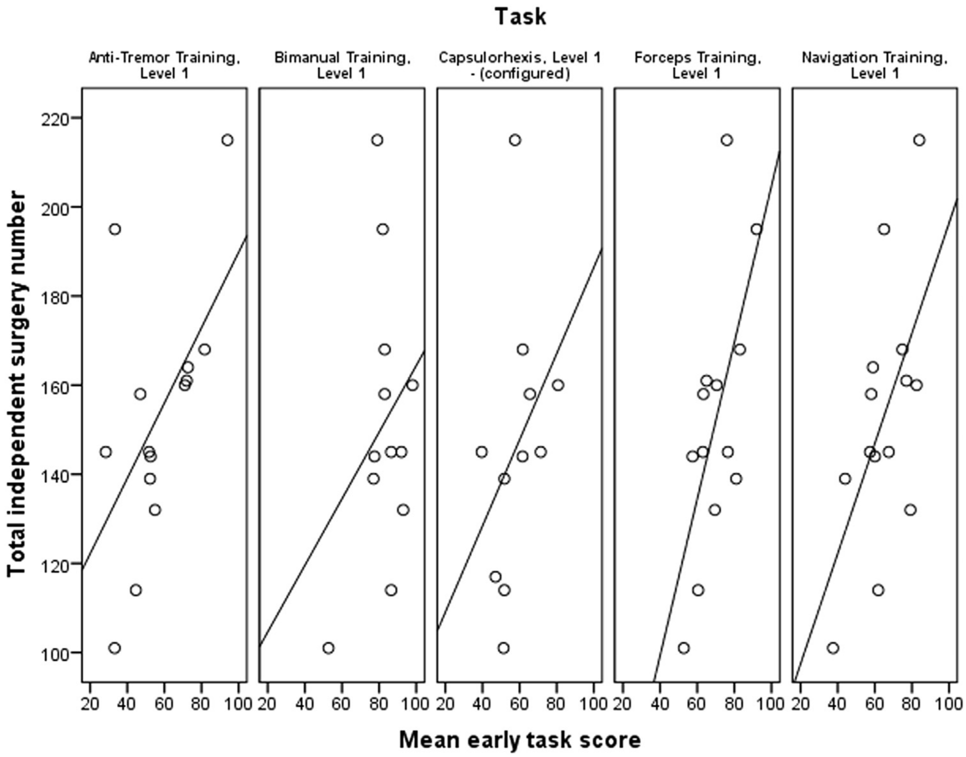

FIGURE 3. Scatter plot showing relationship between primary number of cataract surgery and each early mean task score.

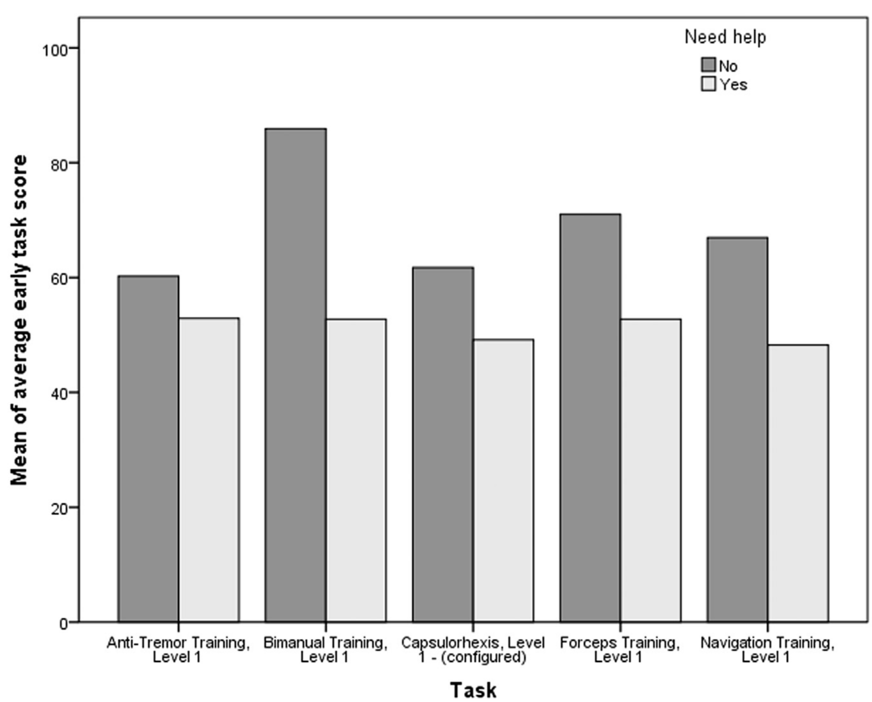

FIGURE 4. Bar chart showing relationship between task scores among residents needing extra help during surgery and those who did not.

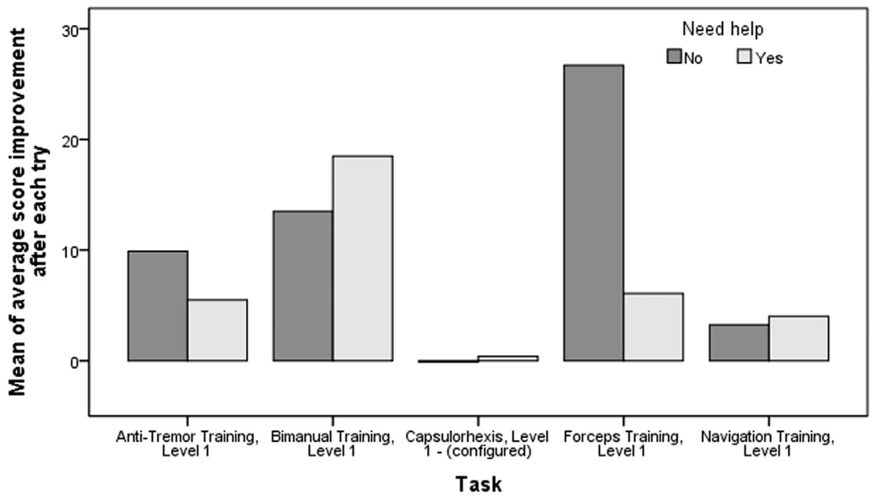

FIGURE 5. Bar chart showing improvement in task score after each try in all residents needing extra help or not.

OR performance characteristic corresponding to each increase of 10 in Eyesi score is shown in Table 4.

The ICC of the raters' GRASIS scores was 63%. Comparison of GRASIS score in residents needing extra help versus others is shown in Table 5 and the scatter plot in Figure 6.

DISCUSSION

To our knowledge, this is the first study identifying a link between ophthalmology resident surgical performance and early surgical simulator scores. We found a statistically significant correlation between Eyesi capsulorhexis score and total cataract surgeries done by residents. The capsulorhexis module is considered by many cataract surgeons to be the most accurate model of live surgery within the available Eyesi modules,[17] strengthening our confidence in this apparent relationship. We also found significant correlation between antitremor training level 1 and forceps training level 1 on the Eyesi and primary surgery numbers.

We used the GRASIS for assessment of surgical performance as it has been shown to have face and content validity.[15] Although all residents were assessed with the GRASIS, the timing of the assessments varied according to when the residents were on a specific cataract rotation during their senior year. The ICC of the raters was found to be 63%, which is acceptable. We found significant correlation between the mean of total GRASIS score and the mean of Eyesi total task score.

We used the number of primary cataract surgeries because it is an objective measure. There are, however, many confounders such as resident effort, faculty vacations, and seasonal variation that can affect total number of cases independent of surgical skill.

We compared the Eyesi results and OR performance among residents identified by the program director as needing extra surgical training to the remaining group of residents. We found higher Eyesi task scores in residents who did not need extra help during their residency.

Of note, improvement in capsulorhexis level 1 and navigation training level 1 scores was very minimal in both residents who needed extra help during residency and those who did not. This was a surprising finding, though the mean scores for these tasks were much lower in residents needing help during residency.

To date, selection for most surgical training programs is based on assessment of attributes such as intelligence, knowledge, and academic achievements, in addition to an interview process.[18] This classic selection process of candidates is not always sufficient to judge surgical aptitude.[19] Usually technical skills and fundamental attributes that are important to surgical skills are not assessed. This is in contrast to selection methods in military aviation, for example, where such variables are measured and used.[18] Some surgical residency programs consider skill examinations for their applicants, evaluating psychomotor skills, perceptual ability, and visuospatial ability[18]; however, this program is not widespread.

Previous studies distinguish innate surgical ability from surgical skills. Skills are learned but require underlying fundamental abilities, not all of which are known.[18] Abilities such as strength, manual dexterity, hand-eye coordination, psychomotor skills, perceptual ability, and visuospatial ability (including stereopsis) have been identified by various authors as being important for surgical training.[18,20-23] It is not clear, which of these abilities (and what degree of these abilities) are required to become a good surgeon. Given the variety of measures used in other studies, and the varied correlations between measures and performance, there may be well different abilities required

TABLE 4. Mean Change in Operating Room Performance Characteristic Corresponding to Each Increase of 10 in Eyesi Score

Explanatory Score	Total Surgery	Independent Surgery	Instrument Handling	Treatment of Ocular Structure	Treatment of Ocular Structure 1	Time and Motion	Use of Nondominant Hand	Overall Performance	Sum Grasis Score	Odds of No Help Needed
Antitremor training level 1	11.0	8.5	0.1	0.1	0.1	0.1	0.0	0.1	0.8	1.02
Bimanual training level 1	17.7	7.5	0.4	0.4	0.4	0.2	0.3	0.2	3.4	1.22
Capsulorhexis level 1 (configured)	29.9	9.7	0.0	−0.1	−0.1	0.1	−0.2	−0.1	0.2	1.14
Forceps training level 1	14.0	17.6	0.4	0.4	0.4	0.3	0.4	0.3	2.7	1.12
Navigation training level 1	14.8	12.4	0.3	0.3	0.3	0.3	0.3	0.1	2.4	1.14

in different surgical specialties. Although it seems clear that individual differences in ability may influence how rapidly a trainee may progress, we generally do not know what the limits are beyond which a trainee may never become adept in spite of how hard and long they practice.[18] In one study of 20 medical students, Alvand et al.[24] concluded that there is variation in innate arthroscopic skill among future surgeons, with some individuals unable to achieve competence in basic arthroscopic tasks despite sustained practice.

Our study has a number of limitations. It is retrospective with data from 1 residency program. The timing of first use of the Eyesi simulator varied among residents, as did the total time of use and total number of attempts. The decision to exclude the first 3 attempts on the Eyesi in an effort to eliminate familiarity with the instrument as a variable was arbitrary. We used the next 8 attempts as all the included subjects had at least these number of attempts. We used the composite scores on the Eyesi modules that were developed by the manufacturer. It is possible that using raw data rather than composite scores would yield different results. A larger multicenter prospective study with a standard simulator use protocol is needed to confirm our results. Such a study would also benefit from masked review of recorded cataract surgery cases by trained evaluators. We also do not know if analogous tests would be useful in other surgical fields.

In many ophthalmology residencies, cataract surgery is predominantly done in the last year of training, thus, program directors may not know until late in training which residents need additional help and experience to achieve competence. It would be helpful to residents and program directors to have early indicators of which residents may need such help. It is possible that composite task scores on the Eyesi simulator may provide such an early indication, which could be used in conjunction with other indicators that program directors already use such as wet laboratory performance and faculty assessment of minor procedure performance. This would give ample of time for additional instruction and practice in the wet laboratory as well as on the simulator and other methods such as an earlier and more gradual stepwise introduction to the surgery.[25]

Though GRASIS for assessment of surgical performance has been shown to have face and content validity,[15] bias may have been interjected into the methodology based upon the timing of the surgical experience in the third year as well as the pairing of the resident and attending.

As there is overlap in task scores between residents who needed extra help and those who did not in our study, we cannot prospectively identify with certainty which future residents may need extra help. These results should, therefore, be used with caution to avoid incorrectly labeling a resident as a poor surgeon. With a larger data set obtained prospectively it may be possible to construct a mathematic model from multiple Eyesi scores that could predict this more accurately and find a threshold for this prediction

TABLE 5. Comparison of Grasis Score in Residents Needing Extra Help Versus None

	Help		
	No	Yes	p*
Use of instrument handling	**4.47** ± **0.51**	**3.42** ± **0.48**	**0.002**
Treatment of ocular structure	4.45 ± 0.53	3.77 ± 0.63	0.037
Time and motion	4.24 ± 0.53	3.27 ± 0.43	0.004
Use of nondominant hand	4.43 ± 0.51	3.79 ± 0.57	0.043
Overall performance	4.39 ± 0.53	3.51 ± 0.44	0.015
Sum	21.94 ± 2.34	17.73 ± 2.35	0.005
Mean	4.39 ± 0.47	3.55 ± 0.47	0.005

*Based on Mann-Whitney test.

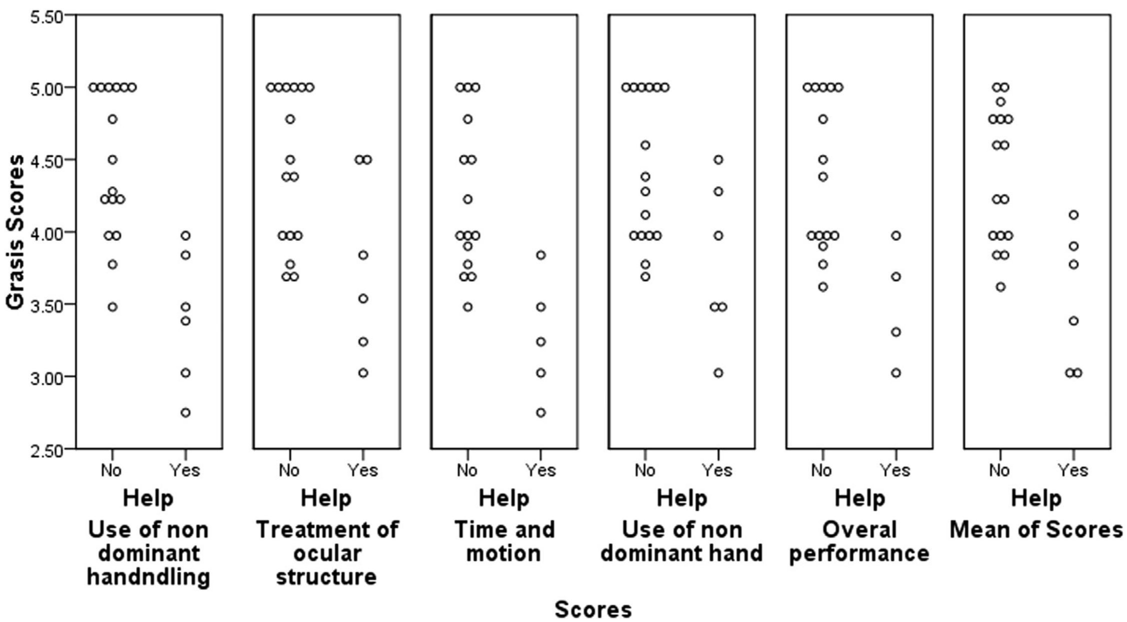

FIGURE 6. Scatter plot showing GRASIS score for residents needing extra help versus those who did not.

A larger data set obtained prospectively might also allow us to approach the question of whether we can assess the surgical aptitude of applicants to ophthalmology programs by use of simulation. Development of frequency distributions of ability scores on tests and use of a composite model could identify aptitude cutoff scores, which can differentiate applicants with less potential from others.[19] This could prevent the unusual but very difficult situation of determining late in training that a resident may have sufficient difficulty in acquiring surgical skills that they cannot become proficient within the 3 years of ophthalmology residency. This is a high-stakes assessment, so this endeavor must be approached with caution.

In summary, we found that eye surgery simulator task scores early in ophthalmology residency correlate with subsequent surgical performance. Early task scores on Eyesi may allow early identification of residents who need more help in learning surgery.

Owing to the limitations of our study, because of its retrospective nature and specific criteria we defined for our study, a prospective trial with a large number of subjects and a defined protocol is recommended.

REFERENCES

1. Nagendran M, Gurusamy KS, Aggarwal R, Loizidou M, Davidson BR. Virtual reality training for surgical trainees in laparoscopic surgery. *Cochrane Database Syst Rev*. 2013;27(8):CD006575. http://dx.doi.org/10.1002/14651858.CD006575.pub3.

2. Walsh CM, Sherlock ME, Ling SC, Carnahan H. Virtual reality simulation training for health professions trainees in gastrointestinal endoscopy.

Cochrane Database Syst Rev. 2012;13(6):CD008237. http://dx.doi.org/10.1002/14651858.CD008237.pub2.

3. Daly MK, Gonzalez E, Siracuse-Lee D, Legutko PA. Efficacy of surgical simulator training versus traditional wet-lab training on operating room performance of ophthalmology residents during the capsulorhexis in cataract surgery. *J Cataract Refract Surg*. 2013;39(11):1734-1741.

4. Kirkman MA, Ahmed M, Albert AF, Wilson MH, Nandi D, Sevdalis N. The use of simulation in neurosurgical education and training. A systematic review. *J Neurosurg*. 2014;121(2):228-246.

5. Sikder S, Tuwairqi K, Al-Kahtani E, et al. Surgical simulators in cataract surgery training. *Br J Ophthalmol*. 2014;98(2):154-158.

6. Thomas GW, Johns BD, Marsh JL, Anderson DD. A review of the role of simulation in developing and assessing orthopaedic surgical skills. *Iowa Orthop J*. 2014;34:181-189.

7. Mahr MA, Hodge DO. Construct validity of anterior segment anti-tremor and forceps surgical simulator training modules. *J Cataract Refract Surg*. 2008;34(6):980-985.

8. Privett B, Greenlee E, Rogers G, Oetting TA. Construct validity of a surgical simulator as a valid model for capsulorhexis training. *J Cataract Refract Surg*. 2010;36(11):1835-1838.

9. Rossi JV, Verma D, Fujii GY, et al. Virtual vitreoretinal surgical simulator as a training tool. *Retina*. 2004;24(2):231-236.

10. Le TD, Adatia FA, Lam WC. Virtual reality ophthalmic surgical simulation as a feasible training and assessment tool: results of a multicentre study. *Can J Ophthalmol*. 2011;46(1):56-60. http://dx.doi.org/10.3129/i10-051.

11. Maan ZN, Maan IN, Darzi AW, Aggarwal R. Systematic review of predictors of surgical performance. *Br J Surg*. 2012;99(12):1610-1621.

12. Lynch J, Aughwane P, Hammond TM. Video games and surgical ability: a literature review. *J Surg Educ*. 2010;67(3):184-189.

13. Kolga Schlickum M, Hedman L, Enochsson L, Kjellin A, Felländer-Tsai L. Transfer of systematic computer game training in surgical novices on performance in virtual reality image guided surgical simulators. *Stud Health Technol Inform*. 2008;132:210-215.

14. Vassiliou MC, Feldman LS. Objective assessment, selection, and certification in surgery. *Surg Oncol*. 2011;20(3):140-145.

15. Cremers SL, Lora AN, Ferrufino-Ponce ZK. Global rating assessment of skills in intraocular surgery (GRASIS). *Ophthalmology*. 2005;112(10):1655-1660.

16. Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics*. 1977;33(1):159-174.

17. Henderson BA, Ali R. Teaching and assessing competence in cataract surgery. *Curr Opin Ophthalmol*. 2007;18(1):27-31.

18. Gallagher AG, Leonard G, Traynor OJ. Role and feasibility of psychomotor and dexterity testing in selection for surgical training. *ANZ J Surg*. 2009;79(3):108-113.

19. Panait L, Larios JM, Brenes RA, et al. Surgical skills assessment of applicants to general surgery residency. *J Surg Res*. 2011;170(2):189-194.

20. Sachdeva R, Traboulsi EI. Performance of patients with deficient stereoacuity on the EYESi microsurgical simulator. *Am J Ophthalmol*. 2011;129(6):781-784.

21. Waqar S, Williams O, Park J, Modi N, Kersey T, Sleep T. Can virtual reality simulation help to determine the importance of stereopsis in intraocular surgery? *Br J Ophthalmol*. 2012;96(5):742-746.

22. Cuschieri A, Francis N, Crosby J, Hanna GB. What do master surgeons think of surgical competence and revalidation? *Am J Surg*. 2001;182(2):110-116.

23. Gallagher AG, Cowie R, Crother I, Jordan-Black JA, Satava RM. PicSOr: an objective test of perceptual skill that predicts laparoscopic technical skill in three initial studies of laparoscopopic performance. *Surg Endosc*. 2003;17(9):1468-1471.

24. Alvand A, Auplish S, Gill H, Rees J. Innate arthroscopic skills in medical students and variation in learning curves. *J Bone Joint Surg Am*. 2011;Dec;93(12):1586-1591.

25. Henderson BA, Kim JY, Golnik KC, et al. Evaluation of the virtual mentor cataract training program. *Ophthalmology*. 2010;117(2):253-258.